An Analytic and Probabilistic Approach to the Problem of Matroid Representibility

D. Capodilupo*[1], S. B. Damelin†[2], S. Freedman‡[1], M. Hua§[1], J. Sun¶[1], and M. Yu‖[3]

[1]Department of Mathematics, University of Michigan
[2]Mathematical Reviews, The American Mathematical Society
[3]Australian National University

Abstract

We introduce various quantities that can be defined for an arbitrary matroid, and show that certain conditions on these quantities imply that a matroid is not representable over $\mathbb{F}_q$. Mostly, for a matroid of rank r, we examine the proportion of size-$(r-k)$ subsets that are dependent, and give bounds, in terms of the cardinality of the matroid and q a prime power, for this proportion, below which the matroid is not representable over $\mathbb{F}_q$. We also explore connections between the defined quantities and demonstrate that they can be used to prove that random matrices have high proportions of subsets of columns independent.

1 Introduction of Quantities

By a subset of a matrix we will mean a subset of its columns, and by its size we will mean the total number of columns it has. We will say a matroid is q-representable if it has a matrix representation over $\mathbb{F}_q$.

1.1 A Generalization of Uniformity

First, we generalize a basic definition.

Definition 1. *A matroid M of rank r is said to be **uniform** if every size-r subset of M is independent.*

Definition 2. *We define the k-**dependence** of a matroid of rank r as the proportion of its size-$(r-k)$ subsets that are dependent. When a matrix has k-dependence 0, we call it k-**independent**, otherwise we call it k-**dependent**. For a matroid M, we will denote rank by $r(M)$, cardinality by $s(M)$, and k-dependence by $d(M,k)$.*

Note that, by these definitions, a matroid M is uniform if $d(M,0)=0$, i.e., if it is 0-independent.

1.2 Optimal Representable Matrices

It is natural to try to optimize some property of a matroid given given certain constraints, especially q-representability. We use the following symbols to denote optimal achievable quantities:

Definition 3.

*dcapo@umich.edu
†damelin@umich.edu
‡samjfree@umich.edu
§mikwa@umich.edu
¶jeffjeff@umich.edu
‖ming.yu@anu.edu.au

- *By* $Ind_q(r,k,d)$*, we mean the largest* s *such that there exists some full-rank* $r \times s$ *matrix* M *over* $\mathbb{F}$ *with* k*-dependence* $\leq d$*. Equivalently, it is the size of the largest* q*-representable rank-*r *matroid with* k*-dependence* $\leq d$*.*
- *By* $D_q(r,k,s)$*, we mean the smallest* d *such that there exists some full-rank* $r \times s$ *matrix* M *over* $\mathbb{F}_q$ *with* k*-dependence* $\leq d$*. Equivalently, it is the smallest* k*-dependence of any* q*-representable rank-*r *matroid of size* s*.*

These quantities prove useful because we can use them to say the following:

Lemma 1. *Let* M *be a matroid. If, for some* k*,*

- *If* $Ind_q(r(M),k,d(M,k)) \leq s(M)$ *or*
- *If* $D_q(r(M),k,s(M)) \leq d(M,k)$ *for some* k*,*

then M *is not* q*-representable.*

2 Equivalences of Bounds

An equivalence between bounds on Ind and on D exist due to the following:

Lemma 2. *As a function of* s*,* $D_q(r,k,s)$ *is non-decreasing.*

Proof. Let M be a minimally k-dependent q-representable matroid of size s. That is, because we are dealing with finite sets and infima are always achievable,

$$d(M) = D_q(r(M),k,s(M)).$$

Then, for every matroid M' obtained by deletion of one element from M,

$$d(M') \geq D_q(r(M'),k,s(M')) = D_q(r(M)-1,k,s(M)-1).$$

Thus, because each size-$(n-k)$ subset is counted an equal number of times in the measurement of the $d(M')$,

$$D_q(r(M),k,s(M)) \geq D_q(r(M)-1,k,s(M)-1).$$

□

The equivalence between bounds can be stated thus:

Lemma 3.

- *If, for some* $q,r,k,d,Ind_q(r,k,d) < s$*, then, for any* $s' \geq s$*, it holds that*

$$D_q(r,k,s') > d.$$

- *If, for some* $q,r,k,s,D_q(r,k,s) > d$*, then, for any* $d' \leq d$*, it holds that*

$$Ind_q(r,k,d') < s.$$

3 Explicit Bounds

We give various explicit bounds on Ind and D, on whichever of the two the explanation of the bound is simplest. In each case, the equivalent statement on the other function is implied.

Theorem 1. $Ind_q(r,k,0) \leq q^{k+1}(r-k-1)$

Proof. Suppose some matroid M is representable q-representable. Then some $r(M) \times s(M)$ matrix M' over $\mathbb{F}_q$ can be constructed with all size-$(n-k)$ subsets independent.

We treat the columns of M' as vectors in $\mathbb{F}_q^r$, and assume that none of them are the zero vector.

Observe that at most $r-k-1$ columns of M' can lie within a $(r-k-1)$-plane. This implies that the proportion between $(r-k-1)$ and the number of points in an $(r-k-1)$-plane bounds the proportion of the total number of vectors in $\mathbb{F}_q^r$ that are represented as columns in M'.

Explicitly, this proportion is

$$\frac{r-k-1}{q^{r-k-1}}$$

out of

$$q^r$$

vectors in the space. Thus, the total number of columns is bounded by

$$q^r \frac{r-k-1}{q^{r-k-1}} = q^{k+1}(r-k-1).$$

□

Theorem 2. *For some integer n, $D_q(r,k,n\frac{q^n-1}{q-1})$ is minimized by the matrix M consisting of n copies of each unique nonzero vector in $\mathbb{F}_q^r$ up to scaling.*

That is, the matrix consists of exactly n representatives of each point in the projective space.

Proof. Let M as above. The claim is clearly true for $k = n-2$, in which case M is the only vector of the required size that is $(n-2)$-dependent. We proceed inductively. Let M as above, $\vec{v} \in M$. We can view M as a multiset of points in the projective space $\mathbb{P}^{r-1}(\mathbb{F}_q)$. Let $\mathbb{P}^{r-2}$ be some hyperplane in $\mathbb{P}^{r-1}(\mathbb{F}_q)$. Then a set S including one copy of $\vec{v}$ is independent if and only if the projection of $S \setminus \{\vec{v}\}$ from $\vec{v}$ onto $\mathbb{P}^{r-2}$ is independent. A set containing two copies of $\vec{v}$ is dependent. Thus, removing all copies of $\vec{v}$ from M, we can count the number of independent size-$(n-k)$ sets containing $\vec{v}$ by counting the number of independent size-$(n-k-1)$ points of the projection of the remaining members of M onto $\mathbb{P}^{r-2}$ as above. If M contains one column for each vector in $\mathbb{F}_q^r$, then exactly $q+1$ vectors will be projected to each point in the hyperplane. The $(k-1)$-dependence for that arrangement of vectors in the hyperplane, by the inductive hypothesis, is optimal. □

4 Random Matrices

Defining two more quantities, this approach can be used to prove that, with very high probability, a very high proportion of the subsets of a certain size of a random matrix are independent.

By "random matrix," we mean a matrix whose columns are randomly chosen nonzero vectors.

Definition 4. *Let $Ind_q(r,k,d,p)$ be the largest s, or $D_q(r,k,s,p)$ the smallest d, or $P_q(r,k,d,s)$ the smallest p, such that, with probability $1-p$, a random $r \times s$ matrix has k-dependence $\leq d$.*

Lemma 4. *Denote the probability that $(r-k)$ nonzero vectors chosen randomly from $\mathbb{F}_q^r$ are independent by $\pi_{q,r,k}$. Then,*

$$\pi_{q,r,k} = \prod_{i=0}^{r-k} \frac{q^r - q^i}{q^r - 1}.$$

Proof. Each term of the product divides the number of points outside an i-plane by the number of nonzero points in the space. This is the probability that, given that we have already picked i independent vectors, that the next one we pick will lie outside the span of those i. □

Theorem 3.

$$D_q(r,k,s) \leq 1 - \pi_{q,r,k}.$$

Proof. Take a certain choice of $(r-k)$ distinct integers between 1 and r. These correspond to a single size-$(r-k)$ subset of a matrix of size s. Then, the proportion of this particular subset of all size-s matrices that are independent is equivalently $\pi_{q,r,k}$. Since this proportion is equal for any choice of subset, we have that the proportion of all size-$(r-k)$ subsets of all matrices of size s is $\pi_{q,r,k}$. Thus, some matrix achieves this proportion. $\square$

Corollary 1. *$1-\pi_{q,r,k}$ is the mean k-dependence of all $r \times s$ matrices without zero columns.*

Because $\pi_{q,r,k}$ is in general very close to one, viewing p as a proportion of the set of all $r \times s$ matrices, we can get bounds on $D_q(r,k,s,p)$. Specifically,

Theorem 4. *For any q, r, k, s, p,*

$$D_q(r,k,s,p) \leq \frac{1-\pi_{q,d,k}}{p}.$$

Corollary 2. *For any q, r, k, s, d,*

$$P_q(r,k,s,d) \leq \frac{1-\pi_{q,d,k}}{d}.$$

Note that these quantities do not depend on s.

Acknowledgment: This work resulted from a Research for Undergraduates Program at the University of Michigan during the summer of 2015. Support from Mathematical Reviews and the National Science Foundation is gratefully acknowledged.

References

[1] S. Ball, *On large subsets of a finite vector space in which every subset of a basis size is a basis*, Journal European Math. Soc, J. Eur. Math. Soc, 2012, 14(3): 733-748.

[2] S. Ball and J. De Beulle, *On sets of vectors of a finite vector space in which every subset of basis size is a basis II*, Des. Codes Cryptogr. 65 (2012), no. 1-2, 514.

[3] S. Ball, C. Padro, Z. Weiner and C. Xing, *On the representibility of the bi-uniform matroid*, arXiv 1407.7283v1.

[4] F. J. McWilliams, N. J .A. Sloane, *The Theory of Error Correcting Codes*, Amsterdam, The Netherlands: North Holland, 1977, pp. 323.

[5] K. A. Bush, *Orthogonal arrays of index unity*, Ann. Math. Statist., 23 (1952) 426434.

[6] J. W. P. Hirschfeld, *Maximum sets in finite projective spaces*, Surveys In Combinatorics, London Math. Soc. Lecture Note Series 82, Cambridge University Press, Cambridge, 1983, pp. 5576.

[7] J. W. P. Hirschfeld and G. Korchmaros,*On the embedding of an arc into a conic in a finite plane*, Finite Fields Appl., 2 (1996) 274292.

[8] J. W. P. Hirschfeld and L. Storme, *The packing problem in statistics, coding theory and finite projective spaces, update 2001*, in Developments in Mathematics, 3, Kluwer Academic Publishers. Finite Geometries, Proceedings of the Fourth Isle of Thorns Conference, pp. 201-246.

[9] M. Hu, D. Capodilupo, S. B. Damelin, S. Freeman, J. Sun and M. Yu, *A Note on Truncated Pascal Coding Matrices and Lower Bounds for the Number of Linearly Independent Vectors of Fixed Length over F_q*, submitted.

[10] M. Hua and J. Sun, *Personal Communication: Math Overflow: Truncation Lemma*, May 2015.

[11] R. Lidl and H. Niederreiter, *Finite Fields*, Sec. Ed., Encyclo. Math. and Appl., Vol. 20, Cambridge Univ. Press, 1997.

[12] J. Oxley, *Matroid Theory*, Oxford University Press, 1992.

[13] B. Segre, *Introduction to Galois geometries*, Atti Accad. Naz. Lincei Mem., 8 (1967) 133-236.

[14] G. Tallini, *Le geometrie di Galois e le loro applicazioni alla statistica e alla teoria dellinformazione*, Rendiconti di Matematica (3-4), 19(1960), pp. 379-400.

[15] J. F. Voloch, *Complete arcs in Galois planes of non-square order*, in: Advances in Finite Geometries and Designs, Oxford University Press, Oxford, 1991, pp. 401-406.

[16] M. Yu, P. Sadeghi and N. Aboutorab, *On Deterministic linear network codes braodcast and its relation to matroid theory*,

[17] M. Yu and S. B. Damelin *Overview of independent number, matroid theory, and network coding*, Notes, January 2015.